DOCUMENT RESUME

ED 157 378 FL 009 529

AUTHOR Labov, William
TITLE Where Does the Linguistic Variable Stop? A Response to Beatriz Lavandera. Working Papers in Sociolinguistics, No. 44.
INSTITUTION Southwest Educational Development Lab., Austin, Tex.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Apr 78
NOTE 23p.; For related document, see FL009517

EDRS PRICE MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS *Descriptive Linguistics; Grammar; Language Research; *Language Variation; *Linguistic Competence; Linguistics; *Linguistic Theory; Phonology; *Semantics; *Sociolinguistics; Structural Analysis; Syntax

ABSTRACT

This paper is a response to Lavandera's question regarding the limits of the study of language variation. Sociolinguistics is characterized by its desire to limit representational meaning much more narrowly than formal linguistics. In addition while formal linguistics views language as species-specific and designed to accomodate logical representations, sociolinguistics views language in the context of common biological inheritance. The proper goal of sociolinguistic theory might be stated as the apportionment of the variance on any sub-section of a linguistic system to the functions of representation, identification, and accomodation, and to predict for any new language the probable distribution of information. Thus, variation studies go beyond grammatical description to explanations of variable constraints which will lead to conclusions about the form of grammar. Although in its early days sociolinguistic analysis developed to study sociolinguistic stratification, the full value of variation analysis has only gradually become clear. In addition to its usefulness in describing phonological variation, variation analysis can be a tool in the description of syntactic and semantic analysis, and variation theory as a whole can be a heuristic device for determining the shape of linguistic theory. (AM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED157378

"PERMISSION TO REPRODUCE THIS MATERIAL BY MICROFICHE ONLY HAS BEEN GRANTED BY

James Lawton

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND THE ERIC SYSTEM CONTRACTORS."

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Working Papers in Sociolinguistics

SOUTHWEST EDUCATIONAL DEVELOPMENT LABORATORY
211 East 7th Street
Austin, Texas 78701

FL009529

The *Working Papers in Sociolinguistics* series is produced by the Southwest Educational Development Laboratory with funds from the National Institute of Education, Department of Health, Education & Welfare. The content of these papers do not necessarily reflect DHEW policies or views.

Richard Bauman
Joel Sherzer
EDITORS
The University of Texas, Austin

James H. Perry
EXECUTIVE DIRECTOR
Southwest Educational Development Laboratory

For further information contact:

OFFICE OF COMMUNICATIONS
SEDL
211 East 7th Street
Austin, Texas 78701

Where Does the Linguistic Variable Stop?

A Response to Beatriz Lavandera

by

William Labov

University of Pennsylvania

Sociolinguistic Working Paper

NUMBER 44

April, 1978

Southwest Educational Development Laboratory

211 East Seventh Street

Austin, Texas

RELEVANCE TO EDUCATIONAL PRACTICE

This paper, by William Labov, is a response to a paper by Beatriz Lavandera, "Where Does the Linguistic Variable Stop?" (Working Papers in Sociolinguistics Number 40). While the paper is concerned with theoretical issues in linguistics--essentially the relationship between variation in language use on the one hand and grammatical structure on the other--there is important relevance for educational practitioners. The essential questions asked in any study of linguistic variation are: what aspects of the language are variable (phonology, syntax, vocabulary, etc.); to what degree is this variation due to grammatical constraints and what are they; and to what extent is this variation due to social constraints and what are they? Labov is concerned with how to determine when variants in a language are identical in meaning and when a meaning difference is involved. Educational practitioners should be concerned with the same issue in the speech of their students. In this paper, as elsewhere in his work, Labov insists on placing probabilistic weights on grammatical rules, because of the "ample evidence that human linguistic competence includes quantitative constraints as well as discrete ones."

Joel Sherzer, Editor
Working Papers Series

WHERE DOES THE LINGUISTIC VARIABLE STOP?

A RESPONSE TO BEATRIZ LAVANDERA

William Labov

University of Pennsylvania

The questions raised by Beatriz Lavandera in her paper, "Where does the sociolinguistic variable stop?" are well considered and penetrate to a wide range of issues on the analysis of linguistic variation.[1] At the Linguistic Society meeting of December 1977, where Lavandera's paper was delivered, there was also a remarkable variety of papers that dealt with variation in language. I did not hear any which studied variation for its own sake; each author subordinated his method to a well-defined problem of tracing language change or analyzing language structure. Lavandera has put to us the general question: what are the limits of this technique? How does the study of variation fit into our larger goal of providing an integrated description and explanation of human language?

Linguistic variables or variable rules are not in themselves a "theory of language." They are all heuristic devices. But it is not accidental that linguistic theory has profited from the analysis of variable ways of saying the same thing. Powerful methods of proof proceed from quantitative studies, and this fact is itself a significant datum for our understanding of language structure and language function. Sociolinguistic analysis is normally and naturally associated with a broader view of the use of language than an introspective approach.

No matter how far we penetrate into the details of linguistic structure, our methods of gathering data inevitably return us to the first issue that Lavandera wisely referred to, at the outset. Sociolinguistic analysis asks, "Why does anyone say anything?"

The answer most often given is "to communicate." But that is not very revealing. To communicate what kinds of information?

Though formal linguistics recognizes the existence of expressive and affective information, these are in practice subordinated to what Buhler (1934) called "representational meaning" or what I will call "states of affairs." To be more precise, I would like to say that two utterances that refer to the same state of affairs have the same truth-value, and follow Weinreich in limiting the use of "meaning" to this sense.

Bloomfield's (1926) fundamental postulate of linguistics--essentially that some utterances are partially alike in form and meaning--refers to this truth-conditional sense of "alike" or "same." The sociolinguistic approach hews closely to this line of thinking. Instead of extending meaning as Lavandera suggests, we want to limit it much more narrowly than a formal linguist will do.

Why? The answer is clear when we consider the simple demands of the everyday use of language. How do we know that someone talks like a countryman unless we know that there are rural forms and urban forms with the same meaning? How do we know that someone has spoken politely to us, unless we know that he chose one of several ways of saying the same thing, in this case the more mitigating variant. The two examples refer us to the two major functions of language that are opposed to the representational use:

the self-identification of the speaker, and his accommodation to the listener. To the extent that we recognize their importance, we will take a narrow view of representational meaning.

The formal linguist does just the opposite. He deals with what he knows about--subtle differences in representational meaning. He is programmed to find a difference in meaning between John ate and What John did was eat or between They broke into the liquor closet and The liquor closet was broken into. We see in action two opposing drives: the formalist to expand representational meaning, the sociolinguist to constrict it.

Thus Wolfram as a sociolinguist argues against a host of formal arguments that a-prefixing in a-goin' has no representational meaning (to appear). Sankoff and Thibault (1977) argue against all previous opinion that the auxiliaries avoir and etre have the same meaning in the passe compose. Weiner and Labor (1977) argue that the agentless passive and actives with generalized pronoun subjects say the same thing. Lavandera shows her alertness to the formal viewpoint by refusing to be convinced by these arguments.

It is good that there should be these two opposing imperialisms. The sociolinguist, intent on social variation, might miss some of the subtle differentiations that grammar can make. The formal linguist, insensitive to social variation, may create differences that are as idiosyncratic as the New York City woman who said to me, "The little ones are my [velzɨz]; the big ones are my [vazɨz]."

Along with this goes an opposing view of the place of human language in the larger biological spectrum. The sociolinguist sees language built

on a common biological inheritance. It follows that logical representations are constructed with faculties that were originated to communicate signals of territoriality and accommodation. Therefore the variable component and its characteristic functions are still quite prominent. The formal linguist tends to view human language as a new species-specific creation designed to accept an input of logical representations. In this view, the discrete nature of language categories is their overwhelming characteristic and any variation remaining has little significance.

Can we now rise beyond ideology and construct a higher level theory that gives proper weight to both aspects of human language? This might indeed be described as the proper goal of a sociolinguistic theory: to apportion the variance in any sub-section of a linguistic system to the functions of representation, identification, and accommodation, and to predict for any new language the probably distribution of the information conveyed in prosody, the vowel system, the quantifier system, and so on.

Let us consider what would be the empirical foundations of such a theory.

First, Lavandera argues that sociolinguistic phonological variables convey information on sex, class, etc. We have good evidence on subjective reaction tests that this is the case (Labov 1966:Ch. 11, Labov 1972:247-251). But we should still bear in mind that these tests are quantitative responses to qualitative input. The exemplars of the variables are polar opposites or (at the most) three-valued. No one has done the experiments necessary to prove that differential frequencies convey differential social significance.

In many syntactic variables, there is no problem whatsoever in establishing sameness of representational meaning. Thus negative concord, which has played such a large part in sociolinguistic work, is by definition multiple negation with the same truth value as single negation.

But in the several cases that Lavandera points to there is a real problem. Weiner and Labov argue that They broke into the liquor closet means the same as The liquor closet was broken into. She is right in not being convinced by our arguments. They are persuasive. But we are not in the business of being persuasive: our enterprise demands conclusive demonstration.

How then to proceed?

The first step is to recognize the basis for the technical procedures in the analysis of variation which began with the isolation and definition of the elements that vary along the same dimensions in response to the same state of affairs. In the approach to negative concord, one of the crucial steps was to recognize that negative attraction to first position is distinct from negative concord (Labov 1972b:Ch. 3) though it may be operating in the same sentence. Similarly, in the approach to the factors operating on the choice of passive, it is important to define the envelope of linguistic choices that confront a speaker who is responding to a given state of affairs. Lavandera followed a similar technique in her approach to the choice of tenses in Spanish si-clauses (1975). But she seems to be thinking of some more global, undifferentiated approach to variation in her criticisms that variable rule analyses use "laboriously defined contexts:"

> A characteristic of this kind of study of syntactic variation is that the definition of the variable requires a series of preliminary steps directed at eliminating all the contexts in which the two alternant forms *contrast*, i.e. do not say the same thing. For instance, once Labov and Weiner have decided to examine together the agentless passive of the form *The closet was broken into* and the active sentence *Somebody broke into the closet*, they specify that they are considering only those cases where *somebody* is [-specific]. Cases with *somebody* [+specific] have been pointed out to differ in meaning from the passives. The same applies to all the [+specific] pronouns *they*, *you*, etc., which are left out of the variable.

This is precisely what the business of sociolinguistic analysis is about: it may take a whole year of study and analysis before we can isolate the context where the relevant variation is to be found, gradually isolating those cases where the same formal item has a different linguistic function and setting aside environments where the variation is neutralized or where the rule is categorized. To return to the widest possible defining environment, as she suggests, would reverse the process and lose the precision of analysis we aim at. Another procedure she cites from our passive paper is a good example. We excluded sentential subjects which were extraposed, as in *It is said that John is here*. Extraposition was 100% in spontaneous speech. There were no sentences of the form *That John is here is said*. It would confuse the issue to include

the contrast of They say that John is here with It is said that John is here. In both cases the subject is postposed and so the effects of given vs. new or parallel structure would operate identically. The explanation of either constraint is dependent on their relation to preposing or postposing subject. Since our interest is in the general relation of semantics to syntax rather than a description of the passive, we would lose our sense of direction by throwing every passive into the hopper.

The investigation of variability in the passive is not a study of the passive for its own sake. Our major concern is to measure the relative influence of semantic information--"given vs. new"--and syntactic structure--"parallel subjects"--on the choice of agentless passive vs. generalized pronouns with active sentences. We find that parallel syntactic structure predominates over the "given" vs. "new" effect. Variable rule analyses show that the agentless passive is strongly favored when the subject of the preceding clause is coreferent to the underlying object of the clause under consideration, and that this effect is even stronger if there is a string of two such subjects preceding.

Let us suppose that the generalized pronouns "they" or "you" retain some part of the referential meaning of ungeneralized pronouns. In what way could this account for our findings? It would mean that people used "they" and "you" because they wanted to refer to some particular third or second person subjects, a reference that would be lost if the agentless passive were chosen. We would also have to suppose that this desire became weaker when the preceding subject was coreferential to the object of the

agentless passive. This proposition is a subtle one, but it is the type of issue that recurs continually in our efforts to use the observable features of language to come to a firm conclusion about unobservables.

Lavandera is quite right in saying that we must not avoid the study of differences in meaning. In line with our general program of apportioning variance in linguistic choice among meaning and the various kinds of social significance, we will often encounter linguistic contrasts which potentially distinguish states of affairs but normally serve as social variants. Lavandera's own example of "wiped out" vs. "exhausted" reminds us that this is universally the case with lexical choice. We can prove that there are no true synonyms, in an absolute sense. But stylistic demands force us to substitute one word for another in speech and writing, so that in any given sequence of sentences we use many words as stylistic variants, though each has the potential ability to distinguish particular states of affairs.

The observation of speech events will not be enough to show when a given variant carries a different meaning. Our methods are precise and reliable when our dependent variable is the choice of two forms for the same meaning. When we are dealing with a choice of two meanings for one form it is a different matter. Each case involves inference and persuasive argument, from collocation with other meaningful items, general context, and so on. Discussions of the meaning of Black English Vernacular <u>be</u> are typical of the inconclusive results of this area (Stewart 1966, Fasold 1972). At present, we have no clear methods of proving to another analyst that we are right. It is in this area that we must turn to experimental methods.

The investigation of the potential meaning of get vs. be was referred to by Lavandera: the Jay-Walking Experiment. The general model here is to construct a context which is semantically ambiguous, including a question that demands semantic interpretation of the element in question that is equally ambiguous; thus the interpretation that is made can be fixed in a larger context for a considerable period of time.

The Jay-Walking Experiment concerned the problem of whether the get passive has a meaning different from the be passive. As R. Lakoff pointed out, it seems intuitively that there is such a difference in I got/was arrested to prove a point (1971). The issue was approached experimentally by a "one-question traffic survey" on the streets of Philadelphia.

> It's about cops and jay-walkers. This happened in Milwaukee, where it's a big issue. This man came to a corner. The light was against him. There was a cop on the corner. And there was no cars coming. And he crossed the street
>
> and he got arrested.
>
> or
>
> and he was arrested.
>
> or
>
> and he got arrested to test the law.
>
> or
>
> and he was arrested to test the law.
>
> Do you think that was the right thing to do?

One of the four choices of test sentence was used with each subject. The question at the end is ambiguous, since "that" or "the right thing to do" refer to some action with agent unspecified. If the "he" in the previous sentence is analyzed as the agent of an inchoative sentence, then it was the pedestrian's

action that is to be judged--crossing the street and getting himself arrested to test the law. If the "he" in the previous sentence is analyzed as the patient of a passive sentence, then it is the agent of that sentence whose action is to be judged--the cop's action in arresting the man.

Thus an answer of the type, "No, because I cross the street myself that way all the time" indicates an analysis of the test sentence as a passive. An answer of the type, "No, because he was just asking for trouble" indicates an analysis of the test sentence as an inchoative.

Results of the Jay-Walking experiment as reported in Labov 1975 show that got and be are clearly differentiated in the grammars of most speakers when they are followed by a purpose clause. But the simple "got arrested" and "was arrested" forms produced the same range of responses. This is a beginning in the task of specifying where the get auxiliary serves as a social and stylistic variant of be and where its capacity to distinguish states of affairs is realized.

Techniques of this sort are required if we are to demonstrate whether or not there is a residual referential meaning of the generalized pronouns "they" and "you". It is quite possible that residual reference of these terms is a factor inhibiting the choice of the agentless passive, though the variable rule analyzses indicate that it is likely to be a small effect, since a very large part of the variance is now accounted for.

In carrying out such studies, we are obviously not confined to the effect of external factors on linguistic choices. In this particular case of the passive, the effects of social class, ethnicity and age are minimal, and sex has no effect at all. Lavandera's discussion of variability seems to spring from an earlier period in the study of variation when the primary motivation was to discover the social motivation of particular sound changes, and demon-

strate the ordered distribution of linguistic choices across the social spectrum. But since Labov, Cohen, Robins and Lewis 1968, the analysis of variation has been equally concerned with internal constraints on rule-governed behavior, and the light which these constraints throw on cognitive operations and "knowledge" of the grammar.

The study of -t,d deletion is a phonological study. But it reveals grammatical knowledge in the form of grammatical constraints on the phonological process. Thus consonant cluster simplification shows us that speakers recognize past tense boundaries by deleting less often after such a boundary. The case of the ambiguous clusters lost, told, left is even more interesting. Is the final consonant a past tense marker, separated by a derivational + boundary, or is it part of the stem? An analysis of this issue was carried out by Sally Boyd for 83 subjects whose consonant cluster simplification was studied by our project on linguistic change and variation. The age of the subjects ranged from 5 to 72 years. Variable rule analyses showed that the weight contributed by the presence of an ambiguous cluster to the probability of deletion declined steadily with age from 14 to 71. Figure 1 shows these data with the regression line fitted to them. The result is a remarkable one, because this analysis of objective behavior indicates that the form of the grammar changes steadily with age, as the speaker learns more and more about his language. In this case we are dealing with a deeper analysis of derivational morphology: as the boundary in /los+t/ is recognized with greater clarity, the behavior of lost under -t,d deletion moves steadily away from cost and towards tossed. There is no reason to believe that the same deepening analysis does not occur in other areas of derivational morphology.

The value of the analysis of variation as an indicator of underlying grammatical processes is clear. It seems unlikely that this shift of the

ambiguous class towards past-tense status is a communicative device in itself. The fact that this variation reveals a speaker's grammatical knowledge to us does not mean that it reveals anything to other speakers in the course of spontaneous conversation.

The most detailed investigation of syntactic variation that I have engaged in is the study of acquisition of inversion by my daughter Jessie, by my wife and myself. With a data set of 20,000 questions over 2 1/2 years we were able to trace the transition from a set of phrase structure rules to a single integrated transformational rule of inversion (Labov & Labov 1977). At the point where we begin to write a variable rule of inversion, we do not see any sudden increase in social information conveyed. But we do find evidence of a dramatic increase in Jessie's perception of the relation between sentences, and the integration of her grammatical knowledge into a single schema. At an early stage we discover that contraction favors inversion. But this is not a reasonable result: contraction, which occurs at a much later stage in the derivation, is not likely to be a condition on inversion. This with other evidence leads us to reject the variable rule of inversion that we wrote. At a later stage the favoring effect of contractions disappears, and the fit of observation and prediction improves, and we have reason to believe that the transformational rule does characterize Jessie's grammar. Figure 2 shows the steady decline of the effect of contraction on inversion over time, and the simultaneous decline in the chi-square figure which indicates the increasing fit of observation and prediction.

Thus a variable rule analysis is not put forward as a description of the grammar, but a device for finding out about the grammar. Some results support the initial model, others discredit it. We are left with a statement of the degree of objective evidence for a certain abstract grammatical

relation in the rule system used by a given speaker or a speech community.

It is true enough that our early studies developed as a way of studying sociolinguistic stratification. We didn't realize how massive was the effect of variation in linguistic rules and we have only gradually come to realize the full value of the analysis of variation. It is obvious that Lavandera is correct in saying that the result of an analysis of syntactic variation is not in itself an interpretable finding. It is the explanation of the variable constraints that lead us to conclusions about the form of the grammar. When we reach these conclusions we will not hesitate to place probabilistic weights upon our grammatical rules, no matter where they occur. There is ample evidence that human linguistic competence includes quantitative constraints as well as discrete ones, and that the recognition of such constraints will allow us to build our grammatical theory on the evidence of production and perception in every-day life.

FOOTNOTES

1. This paper was originally given at a Symposium at the December 1977 meeting of the Linguistic Society of America in Chicago, in response to the paper by Beatriz Lavandera with the same title. I am indebted to E. Judith Weiner for emendations at a number of points, as well as for her help in the joint paper which is one of the main subjects of discussion.

REFERENCES

Bloomfield, Leonard. 1926. "A set of postulates for the science of language." LANGUAGE 2:153-164.

Buhler, Karl. 1934. SPRACHTHEORIE. Jena.

Fasold, Ralph. 1972. TENSE MARKING IN BLACK ENGLISH. Arlington: Center for Applied Linguistics

Labov, William. 1966. THE SOCIAL STRATIFICATION OF ENGLISH IN NEW YORK CITY. Arlington: Center for Applied Linguistics.

----------. 1972a. SOCIOLINGUISTIC PATTERNS. Philadelphia: University of Pennsylvania Press.

----------. 1972b. LANGUAGE IN THE INNER CITY. Philadelphia: University of Pennsylvania Press.

----------. 1975. "What is a linguistic fact?" In R. Austerlitz (Ed.), THE SCOPE OF AMERICAN LINGUISTICS. Pp. 159-196. Lisse, Netherlands: Peter de Ridder Press.

----------, Paul Cohen, Clarence Robins, and John Lewis. 1968. A STUDY OF THE NON-STANDARD ENGLISH OR NEGRO AND PUERTO RICAN SPEAKERS IN NEW YORK CITY. Cooperative Research Report 3288. 2 vols. Philadelphia: U.S. Regional Survey.

----------, and Teresa Labov. 1977. "Learning the syntax of questions." In R. Campbell & P. Smith, RECENT ADVANCES IN THE PSYCHOLOGY OF LANGUAGE. Vol. 4B in the NATO Conference Series, Series III: Human Factors. New York: Plenum.

Lakoff, Robin. 1971. "Passive resistance." CLS 7:149-162.

Lavandera, Beatriz. 1975. Linguistic structure and sociaolinguistic conditioning in the use of verbal endings in si-clauses (Buenos Aires Spanish). Unpublished University of Pennsylvania Ph.D. dissertation.

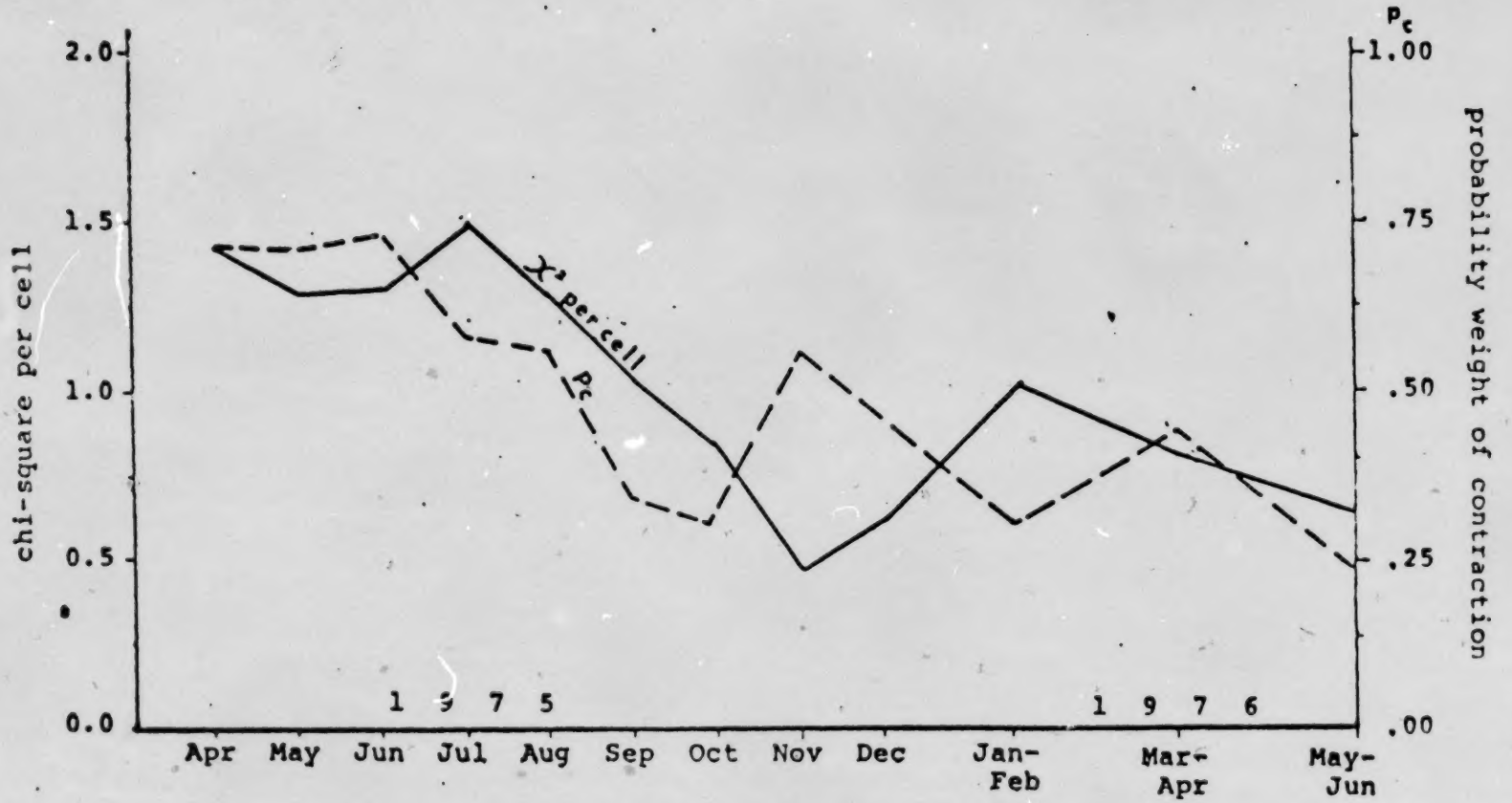
----------. 1977. "Where does the sociolinguistic variable stop?" Paper presented at Linguistic Society of America Meeting, Chicago, December.

Sankoff, Gillian, and P. Thibault. 1977. "L'alternance entre les auxiliaires avoir et etre en francais parle a Montreal." LANGUE FRANCAISE 34:81-108.

Stewart, William A. 1966. "Social dialect, in research planning on language in disadvantaged children." New York: Ferkauf Graduate School, Yeshiva University.

Weiner, E. Judith, and William Labov. July 1977. "Constraints on the agentless passive." Paper presented at Linguistic Society of America Summer Meeting, Honolulu.

Wolfram, Walt. To appear. "A-prefixing in Appalachian English." In W. Labov (Ed.), QUANTITATIVE ANALYSES OF LINGUISTIC STRUCTURE. Academic Press.



Improvement of fit of variable rule model to WH- inversion data compared to decrease in probability factor contributed by contraction

FIGURE 2